



OPEN

Combined transcriptome and metabolome integrated analysis of *Acer mandshuricum* to reveal candidate genes involved in anthocyanin accumulation

Shikai Zhang^{1,2}, Wang Zhan³, Anran Sun^{1,2}, Ying Xie^{1,2}, Zhiming Han^{1,2}, Xibin Qu³, Jiayi Wang³, Laifu Zhang³, Mingshun Tian³, Xuhong Pang³, Jinbao Zhang³ & Xiyang Zhao^{1,2}✉

The red color formation of *Acer mandshuricum* leaves is caused by the accumulation of anthocyanins primarily, but the molecular mechanism researches which underlie anthocyanin biosynthesis in *A. mandshuricum* were still lacking. Therefore, we combined the transcriptome and metabolome and analyzed the regulatory mechanism and accumulation pattern of anthocyanins in three different leaf color states. In our results, 26 anthocyanins were identified. Notably, the metabolite cyanidin 3-O-glucoside was found that significantly correlated with the color formation, was the predominant metabolite in anthocyanin biosynthesis of *A. mandshuricum*. By the way, two key structural genes *ANS* (Cluster-20561.86285) and *BZ1* (Cluster-20561.99238) in anthocyanidin biosynthesis pathway were significantly up-regulated in RL, suggesting that they might enhance accumulation of cyanidin 3-O-glucoside which is their downstream metabolite, and contributed the red formation of *A. mandshuricum* leaves. Additionally, most TFs (e.g., MYBs, bZIPs and bHLHs) were detected differentially expressed in three leaf color stages that could participate in anthocyanin accumulation. This study sheds light on the anthocyanin molecular regulation of anthocyanidin biosynthesis and accumulation underlying the different leaf color change periods in *A. mandshuricum*, and it could provide basic theory and new insight for the leaf color related genetic improvement of *A. mandshuricum*.

Acer mandshuricum, a precious colored-leaf tree in landscapes, belongs to the genus *Acer* (Aceraceae) and mainly distributed among northeastern China¹. Owing to its great ornamental value, *A. mandshuricum* has been selected as the ideal tree species for urban landscaping. It has a visually appealing tree type, and the leaves can change from green to very beautiful red in autumn (Fig. 1)². With the development of the world economy, the process of urbanization is accelerating, and people's demands on the urban environment have been gradually increasing. Obviously, colored-leaf plants can play an important role in the landscape due to their colorful leaves, strong resistance and ease of cultivation. A large number of *A. mandshuricum* trees have been transplanted into urban landscaping³. However, some problems have also appeared. For example, the color change period was greatly shortened when wild *A. mandshuricum* was transplanted to the city. Moreover, it was found that some genes related to colored leaves in natural forest trees could not be passed on steadily to future generations. In view of the important economic and ornamental value of *A. mandshuricum*, it is an urgent problem to study the anthocyanidin accumulation mechanism for color-related genetic improvement.

Numerous scientific studies have shown that changes in the color of leaves are caused by changes in the pigment content of their bodies, including chlorophyll, lutein, carotene, anthocyanin and more^{4–6}. Among the many pigments, anthocyanins have received much attention because of their considerable ability to control leaf color changes⁷. Anthocyanin is a kind of water-soluble natural pigment widely found in plants. There are more than 20 kinds of known anthocyanins, 6 of which are more common in plants: pelargonidin, cyanidin, delphinidin,

¹State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin 150040, China. ²College of Forestry and Grassland, Jilin Agricultural University, Changchun 130118, China. ³Baishishan Forestry Bureau, Jiaohe 132500, China. ✉email: zhaoxyphd@163.com

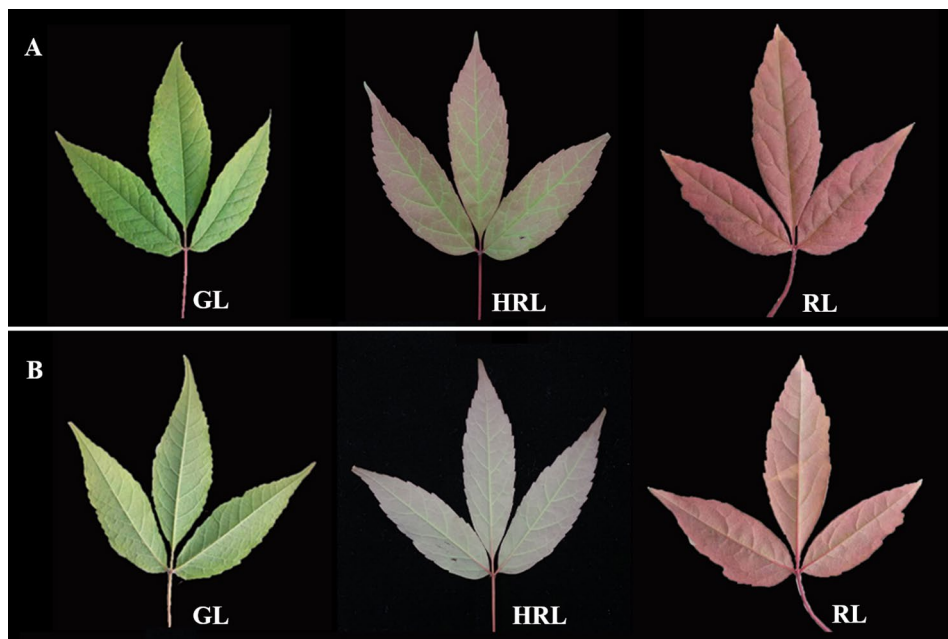


Figure 1. Phenotypes of *A. mandshuricum*. Green (GL), half red leaf (HRL), red leaves (RL) on front (A) and back (B).

peonidin, petunidin and malvidin^{8–10}. The amount of anthocyanins present can directly affect the color change of plants. At present, the synthetic pathway of anthocyanins has been elucidated in many studies, and it is a branch of the phenylpropanoid biosynthetic pathway^{11,12}. The molecular mechanism of key structural genes in its pathways, such as phenylalanine ammonia-lyase (*PAL*), 4-coumarate-CoA ligase (*4CL*), chalcone synthase (*CHS*), chalcone isomerase (*CHI*), dihydroflavonol 4-reductase (*DFR*), anthocyanin synthase (*ANS*) and more, has been clarified^{13–15}. At the same time, the biosynthesis of anthocyanins is also coregulated by some transcription factors, among which the greatest influences were MYBs (v-myb avian myeloblastosis viral oncogene homologs), bHLHs (basic helix-loop-helix proteins), bZIPs (basic leucine zippers) and so on^{16,17}. Therefore, research on the biosynthesis process of anthocyanins in *A. mandshuricum* will help people to better interpret the molecular mechanism of leaf color change and will also contribute to genetic improvement and germplasm innovation.

In recent years, advanced sequencing technology has developed rapidly, and multiomics research has become the main means for people to explore unknown areas of cells^{18,19}. Correlation analysis of the transcriptome and metabolome can realize the co-expression analysis of differential genes and differential metabolites in timing expression^{20–22}. Exploring the causal relationship between genes and metabolites and combining biological function analysis such as functional annotation and enrichment of metabolic pathways is important to identify the key metabolic pathways, key genes and key metabolites^{23–25}. Thus, the relationship between plant regulation mechanisms and biometric function can be analyzed systematically.

At present, the systematic study of leaf color changes and the precise positioning of leaf color-related genes in *A. mandshuricum* have not been reported. In this study, we profiled the transcriptome and metabolome changes in *A. mandshuricum* during various leaf color change stages to investigate the coloration mechanism. Our results shed light on the candidate genes and metabolic pathways underlying the leaf color change in *A. mandshuricum* and could provide fundamental theory and novel insight for the genetic improvement of *A. mandshuricum*.

Results

Metabolome profiling in different color period samples. Throughout the leaf color change cycle of *A. mandshuricum*, it undergoes a transition from green leaves (GL) to half red leaves (HRL) and then to red leaves (RL) (Fig. 1). Therefore, to detect the change in metabolite concentration (mainly focused on anthocyanins) associated with the leaf color change process in *A. mandshuricum*, three types of leaves during different leaf color change periods were used for experiments. We profiled the metabolome of the three samples using widely targeted metabolomics approach. Then, 26 compounds involving anthocyanins were detected and grouped into 8 classes (Table S1, Fig. S1), including 5 cyanidins, 4 pelargonidins, 4 peonidins, 4 malvidins, 4 proanthocyanidins, 2 delphinidins, 2 flavonols and 1 petunidin. Subsequently, the metabolite concentration data were used to perform a hierarchical heatmap clustering analysis of those samples. Among them, most metabolites, such as cyanidins (cyanidin 3,5-O-diglucoside, cyanidin 3-O-(6-O-malonyl-beta-D-glucoside), cyanidin 3-O-arabinoside, cyanidin 3-O-glucoside, cyanidin 3-O-rutinoside), pelargonidins (pelargonidin 3-O-glucoside, pelargonidin 3-O-(6-O-malonyl-beta-D-glucoside) and pelargonidin 3,5-O-diglucoside), peonidins (peonidin 3,5-O-diglucoside and peonidin 3-O-glucoside), proanthocyanidins (procyanidin B1, procyanidin B2, procyanidin B3, procyanidin C1), flavonols (quercetin 3-O-glucoside, fla_dihydrokaempferol) and petunidin 3-O-glu-

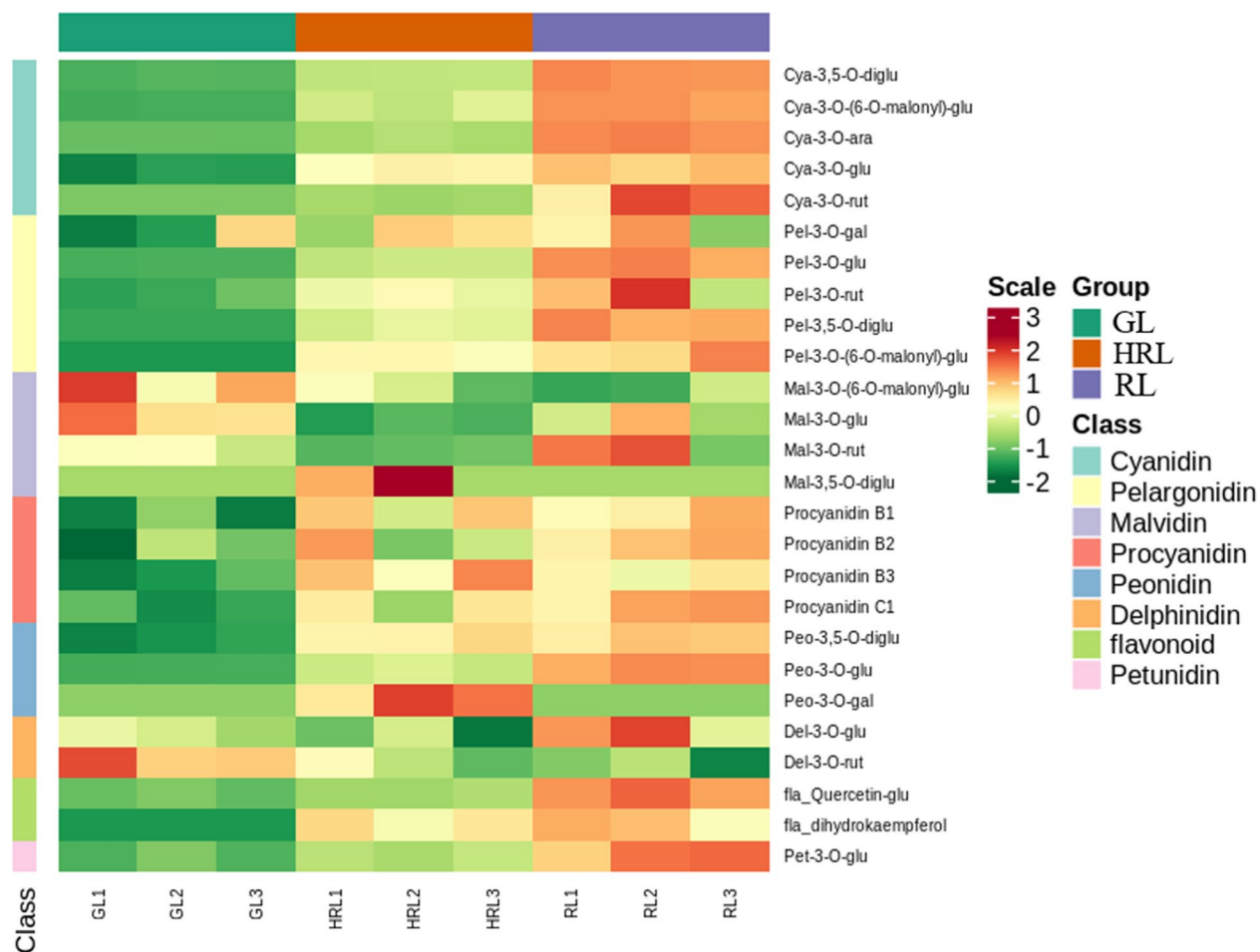


Figure 2. Heatmap of anthocyanidin expression level in different groups.

coside were identified as significantly upregulated during the red leaf period (RL) of *A. mandshuricum* (Fig. 2). From Fig. S2, all the biological duplication groups were successfully grouped together, which indicated that the metabolite group data were highly reliable. In addition, it was obvious that the 9 samples were separated into two classes at the metabolite level, green class (GL) and deep-colored (HRL and RL) samples, and there were great differences in the patterns of metabolite accumulation between the two classes (Fig. S2).

To identify differentially accumulated metabolites (DAMs) between the different samples (GL vs HRL, GL vs HRL_RL, GL vs RL, HRL vs RL) in *A. mandshuricum*, the metabolites with $VIP \geq 1$ and fold change ≥ 2 or ≤ 0.5 were selected, and the heatmap and Venn diagram were used to analyze the differences and expression levels of different metabolites more intuitively and quickly. The results showed that there were significant differences in metabolites between the groups in the three leaf color changes, including 8 (8 up-regulated), 8 (8 up-regulated), 10 (10 up-regulated) and 7 (7 up-regulated) DAMs in GL vs HRL, GL vs HRL, GL vs HRL-RL, GL vs RL, and HRL vs RL, respectively (Table S2). It is worth mentioning that DAMs in the anthocyanin biosynthesis pathway of the four comparative groups were significantly increased (Fig. 3A–D), so we proposed that they were related to the leaf color changes of *A. mandshuricum*, especially cyanidin 3,5-O-diglucoside, cyanidin 3-O-(6-O-malonyl-beta-D-glucoside), cyanidin 3-O-arabinoside, cyanidin 3-O-glucoside, pelargonidin 3-O-glucoside, peonidin 3-O-glucoside, and petunidin 3-O-glucoside. In the Venn diagram (Fig. 3E), 5 anthocyanins were the common metabolites among green-colored and deep-colored samples, and all of them were upregulated and enriched significantly in the anthocyanin biosynthesis pathway (ko00942) (Table S3). Therefore, we speculated that those DAMs produced in the biosynthetic pathways of anthocyanins might be key metabolites during the leaf color change periods of *A. mandshuricum*.

Analysis of transcriptome data. Throughout the leaf color change cycle of *A. mandshuricum*, Illumina HiSeq 2000 was used to construct and sequence nine cDNA libraries based on the total RNAs of green (GL), half-red (HRL) and red (RL) leaf samples of *A. mandshuricum*. A total of 378,613,872 raw reads and 368,399,930 clean reads were obtained, and in the clean reads, the average Q20, Q30 and GC content was 97.63%, 93.42% and 43.89%, respectively (Table S4). Then, gene function annotation was carried out, and we used Blast software²⁶ to compare unigenes to the KEGG (Kyoto Encyclopedia of Genes and Genomes), NR (Non-Redundant Protein Sequence Database), Swiss-Prot (a manually annotated and reviewed protein sequence database), GO (Gene

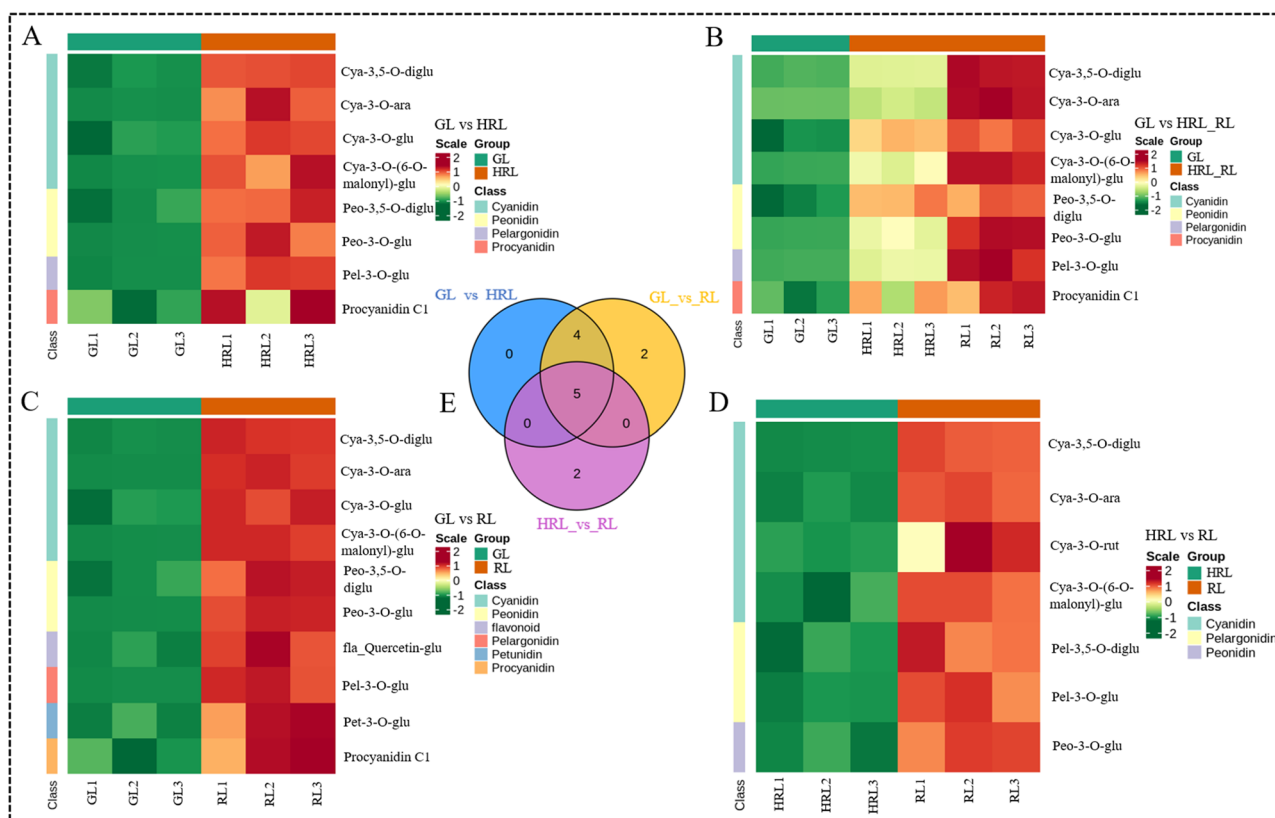


Figure 3. Metabolite expression level analysis of *A. mandshuricum* in different leaf sample groups. (A–D) DAMs in the anthocyanin biosynthesis pathway between four comparative groups of samples of *A. mandshuricum*. These DAMs were analyzed by KEGG enrichment analysis (C, GL vs HRL; D, GL vs HRL_RL; E, GL vs RL; F, HRL vs RL). (E) Venn diagram of metabolites between three compared groups of leaf samples.

Ontology), KOG (euKaryotic Ortholog Groups), Pfam databases (protein family) and Trembl. A total of 201,642 unigenes were successfully annotated, of which 150,946 unigenes (74.86%) were annotated in at least one database above (Table S5 and Fig. S3). Among the 20,162 unigenes, 96,552 (47.88%) unigenes were grouped into 25 KOG classifications; the most highly represented was group R (general function prediction only), and group C (energy production and conversion), group J (translation, ribosomal structure and biogenesis), group O (post-translational modification, protein turnover, chaperones) and T (signal transduction mechanisms) also shared a high percentage of genes (Fig. S4). A total of 116,336 unigenes were categorized into three GO categories (biological process, cellular component and molecular function) and further divided into 60 major functional groups (Fig. S5). To reflect the expression level of transcripts, FPKM was used as an indicator to measure the level of gene expression. Principal component analysis (PCA) of the samples showed that the biological replications of GL samples and RL samples were obviously clustered together and significantly separated from each other, but the three HRL samples were dispersed (Fig. S6). Similar to the metabolomic analysis, there was a significant separation of samples with different colors, indicating that changes in the accumulation of metabolites from green to red leaves were closely controlled by differences in gene expression.

Differentially expressed genes between different colored leaves. DESeq2²⁷ was used to analyze the differential expression between different sample groups, and the results showed that a total of 6988 (3106 down-regulated and 3882 up-regulated), 18,597 (11,442 down-regulated and 7155 up-regulated), and 11,474 (9433 down-regulated and 2041 up-regulated) DEGs were obtained in the GL vs HRL, GL vs RL, and HRL vs RL groups, respectively (Table S6). The differentially expressed genes (DEGs) were identified, and the results are shown in Fig. 4. From the Venn diagram, 217 genes were the common DEGs in all comparison groups, suggesting that those genes across the three comparison groups might be associated with the leaf color change of *A. mandshuricum* (Fig. 4D). Next, we performed GO and KEGG functional annotation and pathway analysis in the three groups and found that the top enriched KEGG terms in the GL vs HRL group were ko01100 (metabolic pathways), ko01110 (biosynthesis of secondary metabolites), ko03010 (ribosome), ko04626 (plant-pathogen interaction) and ko04075 (plant hormone signal transduction) (Fig. S7). Among them, the metabolic pathways and biosynthesis of secondary metabolites were also abundant in the GL vs RL and HRL vs RL groups, indicating that these DEGs might play a key role in leaf color change pathways (Table S7). GO functional enrichment analysis of DEGs between GL vs RL and GL vs HRL revealed that flavonoid metabolic process (GO: 0009812) and flavonoid biosynthetic process (GO:0009813) were significantly enriched (Fig. S8).

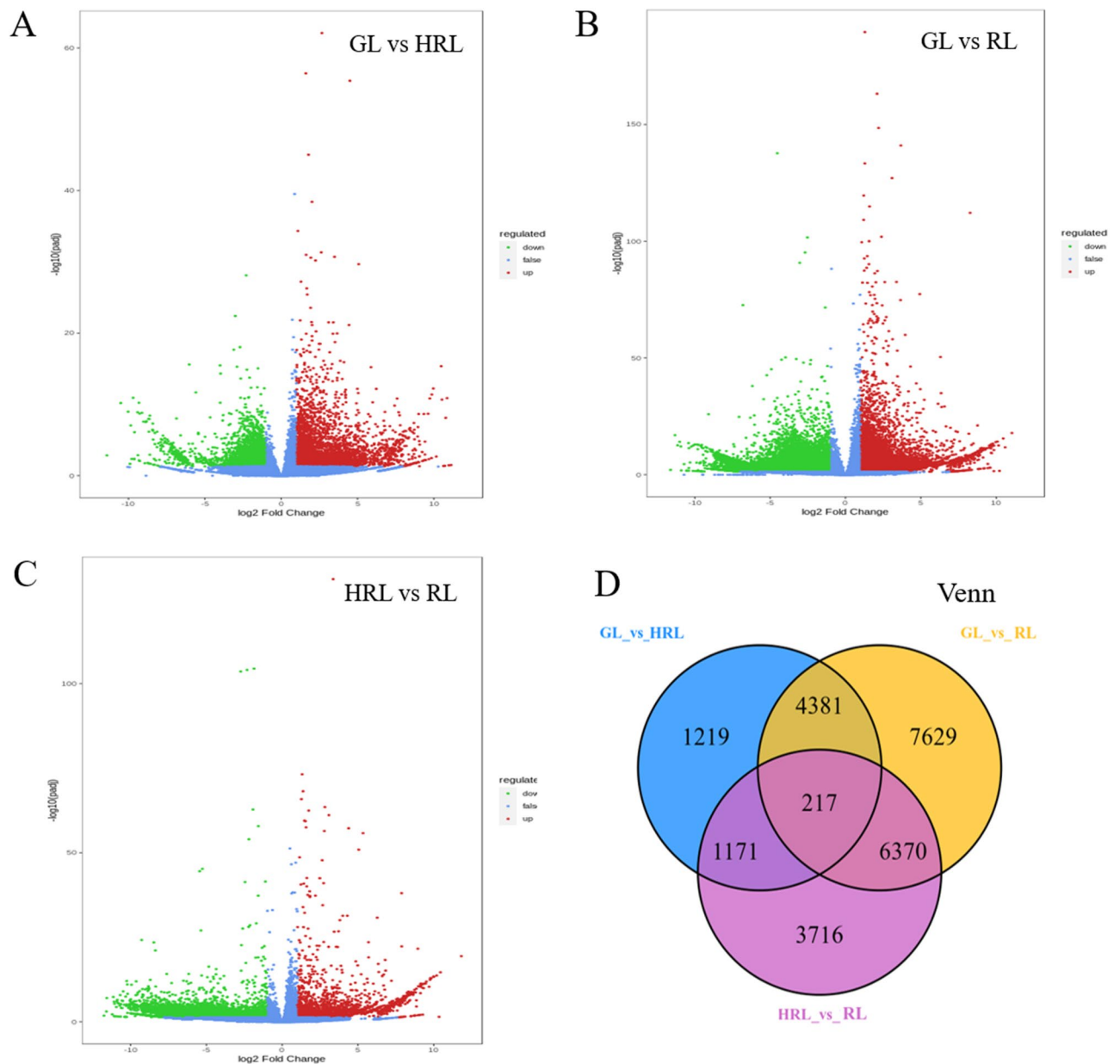


Figure 4. Differentially expressed genes between different colored leaves of *A. mandshuricum*. (A–C) Volcano plots were used to display the down-regulated, no-regulated and up-regulated genes between GL vs HRL, GL vs RL and HRL vs RL. (D) Venn map showing the common and respective DEGs between the three compared groups of leaf samples.

To gain insight into the dynamism of gene expression pattern changes over the leaf color change periods, k-means cluster analysis was performed using the FPKM value of GL, HRL, and RL transcriptome data. Six clusters of co-expressed genes with different expression trends during leaves of different colors were separated by cluster analysis, implying that each cluster may have similar molecular functions (Fig. 5A). Class 1 consisted of 6863 genes, and the expression trend first increased slowly and then decreased rapidly from green to red leaves. By contrast, class 3 contained 3169 genes with an opposite trend compared with class 1. Class 2 encompassed 3479 genes displaying an increase in expression levels from GL to HRL, but a decrease from HRL to RL; this trend was similar to class 4 (1752 genes). A total of 6541 genes were identified in class 5, displaying weak decreases from GL to RL. The most special gene was class 6, which had 2899 genes and exhibited a significant increasing trend in expression level, indicating that they were positively associated with the color change from green to red in *A. mandshuricum*. Several transcription factors (TFs) were encoded by these genes in class 6, predominantly NAC, MYB-related, RWP-RK, GARP-G2-like, bHLH and C2H2, which might play significant roles in the regulation of gene expression levels involved in leaf coloration of *A. mandshuricum* (Fig. 5B).

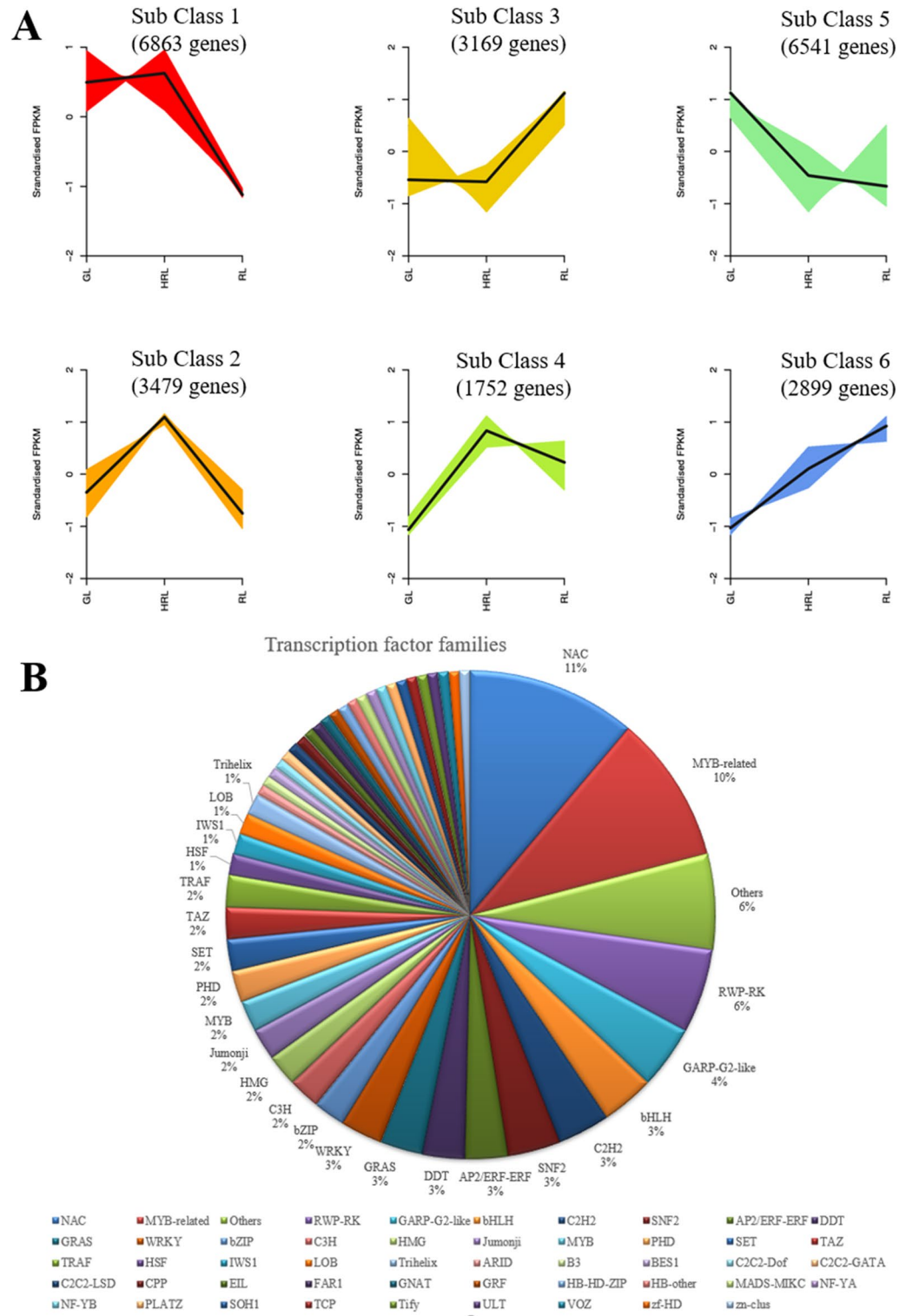


Figure 5. Six classes of co-expressed genes and their kinetic patterns. **(A)** Six classes of co-expressed genes. **(B)** Transcription factors (TFs) were encoded by these genes in class 6.

Modulation of phenylpropanoid, flavonoid, and anthocyanidin biosynthesis pathway genes during leaf color change. Anthocyanin components determine the color of *A. mandshuricum* leaves. However, the key genes involved in *A. mandshuricum* anthocyanin synthesis have not been reported. In this study, we constructed phenylpropanoid, flavonoid, and anthocyanidin biosynthesis pathways in *A. mandshuricum*. Ninety DEGs were searched based on enrichment analysis (Table S8) and predicted the molecular mechanisms leading to red leaf coloration (Fig. 6).

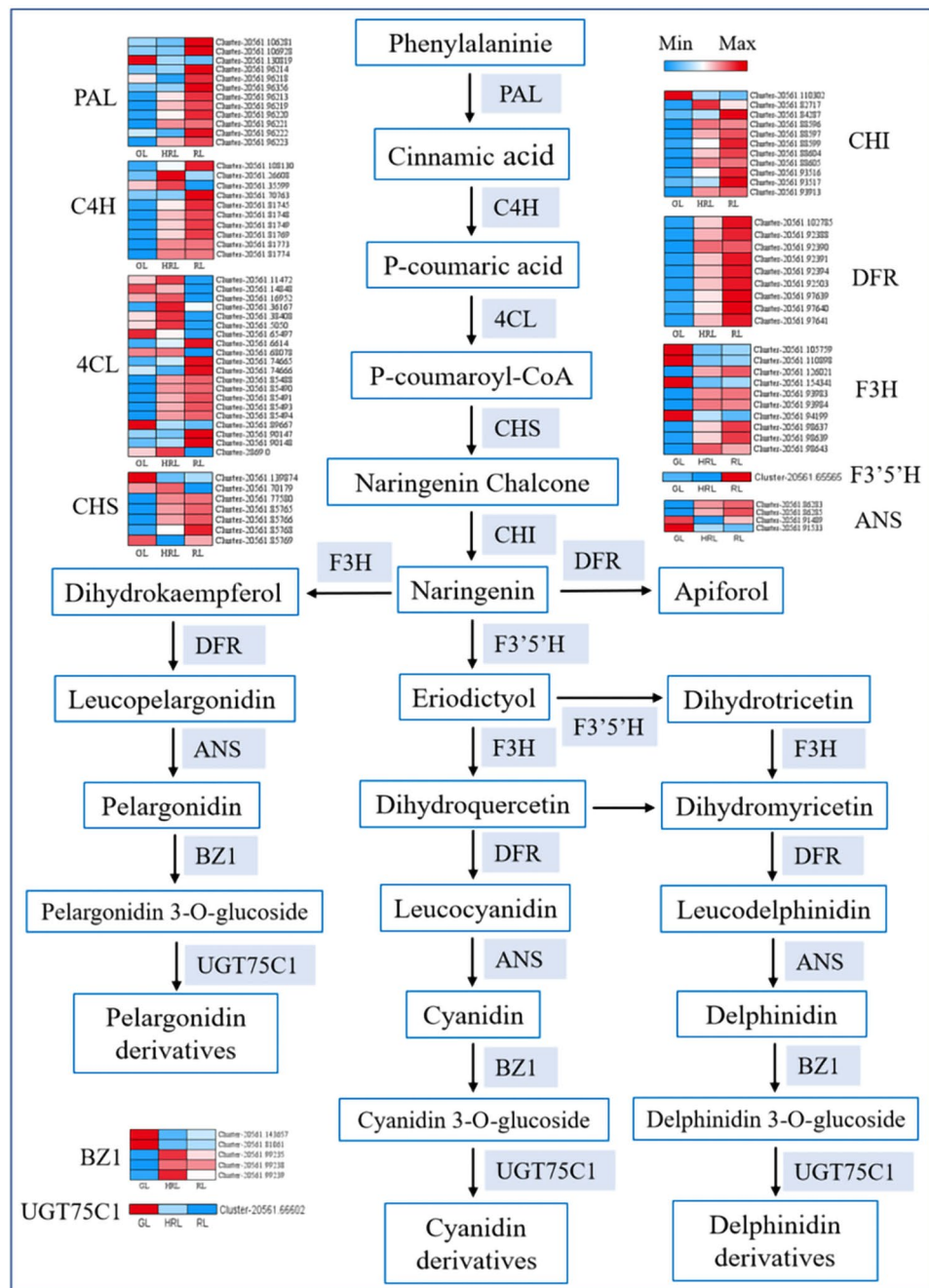


Figure 6. Modulation of phenylpropanoid, flavonoid, and anthocyanidin biosynthesis pathway genes during leaf color change in *A. mandshuricum*. Reconstruction of the phenylpropanoid, flavonoid and anthocyanidin biosynthetic pathways with differentially expressed structural genes. The heat map displays down-regulated and up-regulated structural genes.

From the expression levels, most key pathway genes exhibited significant changes. Indeed, encoding *PAL* (11 DEGs), *C4H* (8 DEGs), *4CL* (10 DEGs), *CHS* (4 DEGs), *CHI* (9 DEGs), *DFR* (9 DEGs), *F3H* (6 DEGs), *F3'5'H* (1 DEG), *ANS* (2 DEGs) and *BZ1* (2 DEGs) enzymes were more highly expressed in HRL and RL than in GL, denoting that those DEGs were the key genes involved in the red pigmentation of *A. mandshuricum* leaves. Conversely, the expression level of one *PAL* gene (*Cluster-20561.130819*) was significantly down-regulated from the phenylalanine to cinnamic acid route when the leaves changed from green to red. Similarly, two *4CL* genes (*Cluster-20561.65497* and *Cluster-20561.89667*), one *CHS* gene (*Cluster-20561.139874*), one *CHI* gene (*Cluster-20561.110302*), four *F3H* genes (*Cluster-20561.105759*, *Cluster-20561.110898*, *Cluster-20564.154341* and *Cluster-20561.94199*), one *ANS* gene (*Cluster-20561.91533*), two *BZ1* genes (*Cluster-20561.143.657* and *Cluster-20561.81061*) and one *UGT75C1* gene (*Cluster-20561.66602*) were also significantly down-regulated in their key pathways, implying that silencing these genes was essential for the leaf change to red. In the meantime,

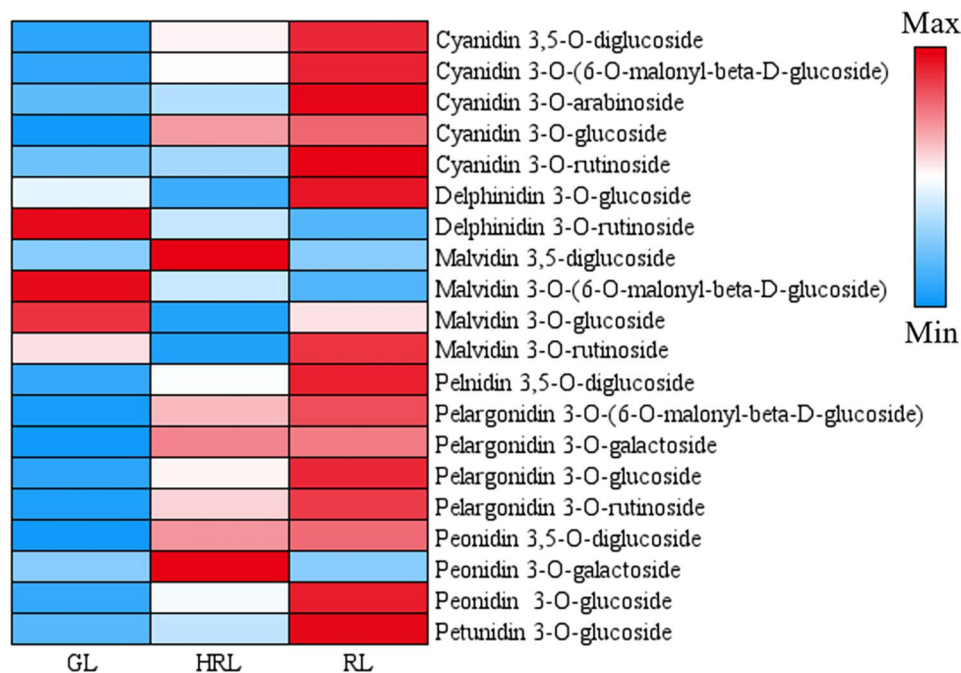


Figure 7. Metabolite accumulation level in anthocyanin synthesis pathways.

some notable genes were also discovered, two *C4H* genes (*Cluster-20561.26608* and *Cluster-20561.35599*), one *4CL* gene (*Cluster-20561.36167*) and one *BZ1* gene (*Cluster-20561.99239*) were activated in HRL stages, but with low expression in GL and RL. The discovery of these key differentially-expressed genes suggested that they could play a pivotal role in the process of *A. mandshuricum* leaf discoloration.

Comparing metabolite accumulation in anthocyanin synthesis pathways (ko00941) (Fig. 7), five cyanidins (up-regulated), two delphinidins (one up-regulated and one down-regulated), four malvidins (two up-regulated and two down-regulated), one pelnidin (up-regulated), four pelargonidins (up-regulated), and four peonidins (up-regulated) were determined. Thus, in summary, our results confirmed that the key genes identified in the biosynthesis pathways of flavonoids and anthocyanins could play a key role in *A. mandshuricum* leaves changing to red. Existing research has confirmed that MYB-related MYB, bHLH, bZIP and other transcription factors (TFs) are widely involved in the regulation of phenyl propane metabolic pathways as regulatory proteins^{28,29}. Analysis of the expression levels of TFs such as bHLH, bZIP and MYB showed that some of them were down-regulated over the red periods, whereas others were significantly up-regulated, implying that these TFs positively regulate flavonoid and anthocyanin biosynthesis pathway genes (Table S9). Overall, these TFs bHLH, bZIP, MYB and more coordinately regulate those key genes, leading to the accumulation of anthocyanin metabolites (cyanidin 3-O-arabinoside, cyanidin 3,5-O-diglucoside, cyanidin 3-O-glucoside, cyanidin 3-O-rutinoside, cyanidin 3-O-(6-O-malonyl-beta-D-glucoside), delphinidin 3-O-glucoside, malvidin 3,5-diglucoside, malvidin 3-O-rutinoside, pelnidin 3,5-O-diglucoside, pelargonidin 3-O-(6-O-malonyl-beta-D-glucoside), pelargonidin 3-O-galactoside, pelargonidin 3-O-glucoside, pelargonidin 3-O-rutinoside, peonidin 3,5-O-diglucoside, peonidin 3-O-galactoside, peonidin 3-O-glucoside, petunidin 3-O-glucoside), giving rise to the beautiful red coloration of *A. mandshuricum* leaves.

Analysis of transcriptome factors. In the biosynthesis process of anthocyanins, it is coregulated by many transcription factors^{30,31}. In this study, we listed anthocyanin biosynthesis-related TFs in *A. mandshuricum*, and a total of 341 (106 up-regulated and 235 down-regulated), 454 (103 up-regulated and 351 down-regulated) and 981 (346 up-regulated and 635 down-regulated) differentially expressed TFs were identified in GL vs HRL, HRL vs RL and GL vs RL, respectively (Table S10). Among them, most represented TFs were annotated as those encoding C2H2, zn-clus, bZIP, MYB-related, NAC, bHLH, C3H, SNF2, TRAF, WRKY, C2C2-GATA and MYB in the GL vs RL group. The most abundant TFs in GL vs HRL (46 up-regulated and 78 down-regulated) were C2H2, NAC, bHLH, MYB-related, C3H, B3-ARE, WRKY and MYB TFs. In the HRL vs RL comparison (9 up-regulated and 151 down-regulated), C2H2, AP2/ERF-ERF, bHLH, bZIP, C2C2-GATA, C3H, HMG, MYB and MYB-related TFs were the most annotated. Previous studies have shown that transcription factors, such as MYB, bHLH and bZIP, have a deeper effect on anthocyanin biosynthesis^{32,33}, and we also found them in many transcription factors in *A. mandshuricum*. For example, most of the MYB TFs (such as *Cluster-20561.80755*, *Cluster-20561.81808*, *Cluster-20561.110291*, *Cluster-20561.85332* and *Cluster-20561.98223*) were co-upregulated with most of the key genes in flavonoid biosynthesis during the discoloration of the leaves in *A. mandshuricum*, and the similar co-upregulated phenomenon was also found in bHLH (*Cluster-20561.117348*, *Cluster-20561.26518*, *Cluster-20561.46232*, *Cluster-20561.92650*, *Cluster-20561.107315*, and more) and bZIP (*Cluster-20561.89129*, *Clus-*

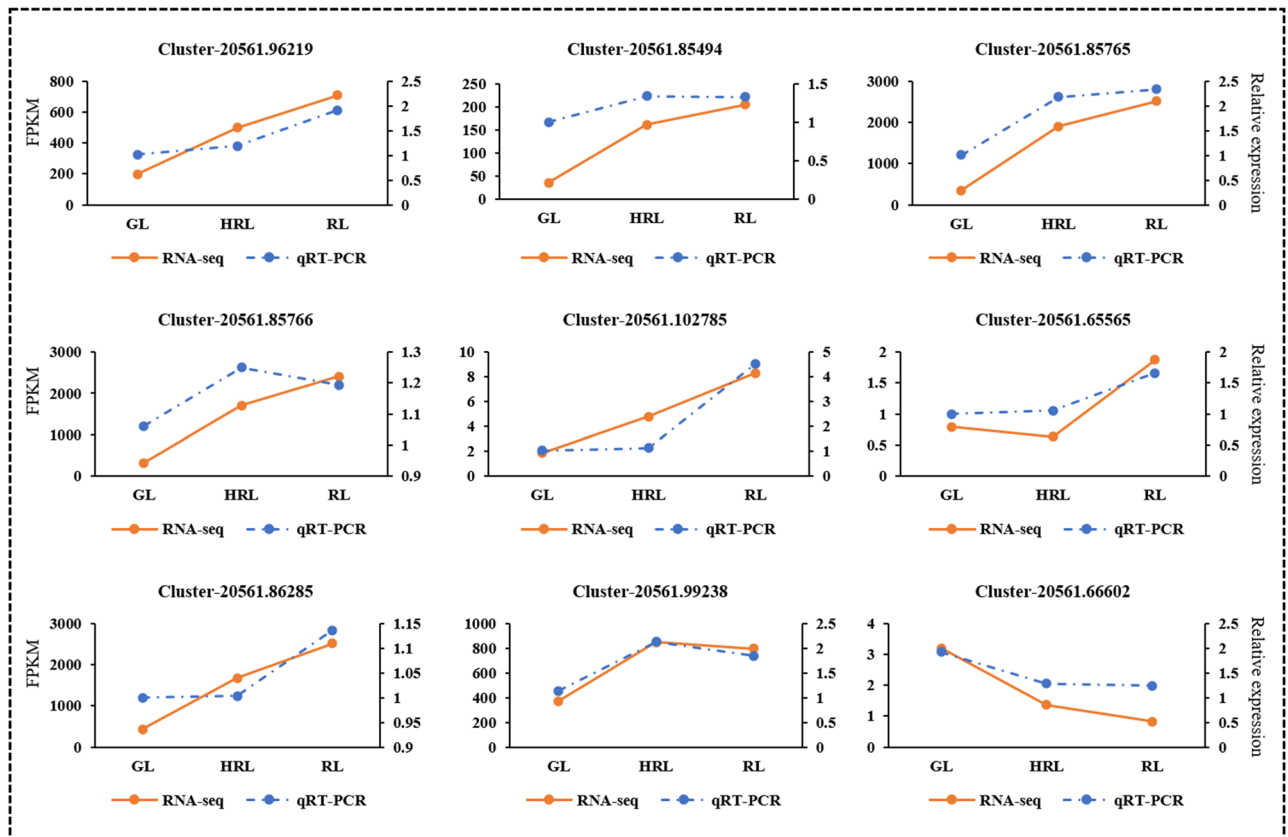


Figure 8. Verification of phenylpropanoid, flavonoid, and anthocyanidin biosynthesis-related differentially expressed genes (DEGs) by quantitative real-time polymerase chain reaction (qRT-PCR).

ter-20561.98504, Cluster-20561.102544, Cluster-20561.141195, Cluster-20561.92791 and so on) TFs, indicating that those MYBs, bHLHs and bZIPs were important positive regulators in leaf color change periods of *A. mandshuricum*. In addition, it is worth noting that in the transcription factors of the three comparison groups, most bHLHs and bZIPs were downregulated when the leaves changed from green to red, indicating that some of the down-regulated genes mediated negative regulation in *A. mandshuricum* anthocyanin biosynthesis (Table S9).

Conjoint analysis between transcriptome and metabolome. To better understand the relationship between genes and metabolites, we mapped the DEGs and DAMs to the KEGG pathmap at the same time. From the results, many DEGs and DAMs were significantly enriched in the anthocyanin biosynthesis pathway. At the same time, we selected those results in which the Pearson correlation coefficient was greater than 0.8. A total of 14,310 DEGs associated with 26 metabolites were found, and most of them jointly regulated the synthesis of one or more metabolites. They were significantly associated with Cya-3,5-O-diglu (6696 genes), Cya-3-O-(6-O-malonyl)-glu (6807 genes), Cya-3-O-ara (6884 genes), Cya-3-O-glu (6568 genes), Cya-3-O-rut (6872 genes), Del-3-O-rut (3789 genes), fla_dihydrokaempferol (6622 genes), fla_Quercetin-glu (6010 genes), Mal-3-O-(6-O-malonyl)-glu (3199 genes), Pel-3,5-O-diglu (6899 genes), Pel-3-O-(6-O-malonyl)-glu (6769 genes), Pel-3-O-glu (6821 genes), Peo-3,5-O-diglu (6755 genes), Peo-3-O-glu (6827 genes), Pet-3-O-glu (6536 genes) and others (Table S11).

From the results of Conjoint analysis between these DEGs and their correlated metabolites, we found two key structural genes *ANS* (Cluster-20561.86285) and *BZ1* (Cluster-20561.99238) in anthocyanidin biosynthesis pathway were significantly up-regulated when the leaves change to red, indicating that they might enhance accumulation of the predominant metabolite cyanidin 3-O-glucoside which is their downstream metabolite, and contributed the red formation of *A. mandshuricum* leaves (Fig. 6).

qRT-PCR validation of DEGs in transcriptome data. To further verify the reliability and accuracy of the transcriptome data, we initially screened 12 DEGs (Cluster-20561.96219, Cluster-20561.85494, Cluster-20561.85765, Cluster-20561.85766, Cluster-20561.102785, Cluster-20561.65565, Cluster-20561.86285, Cluster-20561.99238 and Cluster-20561.66602) which associated with phenylpropanoid, flavonoid, and anthocyanidin biosynthesis to detect expression levels in GL, HRL and RL with qRT-PCR (The primer sequence is in Table S12). The relative expression trends of these candidate genes were similar to those of RNA-seq, indicating the high reliability and accuracy of transcriptome data (Fig. 8). In line with our expectations, the most pivotal genes *ANS* (Cluster-20561.86285) and *BZ1* (Cluster-20561.99238) in anthocyanin biosynthesis of *A. mandshu-*

ricum, were highly expressed when leaves change from green to red, suggesting that they are the candidate key genes underlying the leaf color formation in *A. mandshuricum*.

Discussion

Acer mandshuricum is of vital ornamental and economic value because of its beautiful colored leaves. Previous studies have shown that the main substances in *A. mandshuricum* red leaf extract pigments are anthocyanins and flavonoids². In our present study, using UPLC and a widely targeted metabolomics approach, we investigated the changes in metabolites in three different colored leaf samples of *A. mandshuricum*, aiming to provide a more comprehensive understanding of the accumulation dynamics of anthocyanins in the process of leaf color change. From the results (Table S1), a total of 26 anthocyanidins were simultaneously detected and quantified, mainly cyanidins, pelargonidins, peonidins, malvidins, delphinidins, petunidins, proanthocyanidins and flavonoids, which showed that they played an important role in the critical period of leaf color change. Furthermore, 17 metabolites, including cya-3-O-glu, cya-3,5-O-diglu, pel-3-O-glu, pet-3-O-glu and others, were up-regulated when the leaves changed from green to red. Most of them have been reported in other plant coloration studies^{23,34,35}, indicating that these compounds might be responsible for plants turning red. Especially the metabolite cyanidin 3-O-glucoside was found that significantly correlated with the color formation, was the predominant metabolite in anthocyanin biosynthesis of *A. mandshuricum*, and accumulated in large quantities during the plant red period and was associated with numerous DEGs (6568 genes) ($R^2 > 0.8$) (Tables S1, S10). Therefore, this metabolite, cyanidin 3-O-glucoside, was considered an anthocyanin related to red pigmentation in *A. mandshuricum*. Interestingly, similar results of cyanidin 3-O-glucoside were also reported in some plants with red leaves or fruit^{36–38}.

The biosynthesis of anthocyanin metabolites is tightly controlled by some structural genes (i.e., *PAL*, *CHI*, *DFR*, *ANS* and more), which are also regulated by some transcription factors (i.e., MYB, bHLH, and WD40)^{39,40}. In the whole biosynthesis of anthocyanin metabolites, early biosynthetic genes (*CHS*, *CHI*, *F3H*, etc.) lead to the production of flavonoids, while late biosynthetic genes (i.e., *DFR*, *ANS*) lead to the production of anthocyanins⁴¹. We revealed that most of the structural genes that regulated the phenylpropanoid, flavonoid, and anthocyanidin biosynthetic pathways in the leaf's change to red were highly active (Fig. 6). The expression levels of most genes encoding *PAL*, *C4H*, *4CL*, *CHS*, *CHI*, *F3H*, *DFR*, *F3'5'H*, *ANS* and *BZ1* were higher in RL than in GL, resulting in higher anthocyanin accumulation during the red leaf stage. *F3'H* and *F3'5'H* enzymes are important for anthocyanin accumulation, which could determine the B-ring hydroxylation pattern of anthocyanins⁴², and *ANS* and *DFR* genes have been proven to catalyze the conversion of colorless leucoanthocyanidins into available colored anthocyanidins⁴³. In this study, we found that six *F3H* genes (*Cluster-20561.126021*, *Cluster-20561.93983*, *Cluster-20561.93984*, *Cluster-20561.98637*, *Cluster-20561.98639*, *Cluster-20561.98643*), one *F3'5'H* gene (*Cluster-20561.65565*), nine *DFR* genes (*Cluster-20561.102785*, *Cluster-20561.92388*, *Cluster-20561.92390*, *Cluster-20561.92391*, *Cluster-20561.92394*, *Cluster-20561.92503*, *Cluster-20561.97639*, *Cluster-20561.97640*, *Cluster-20561.97641*) and two *ANS* genes (*Cluster-20561.86283* and *Cluster-20561.86285*) were significantly up-regulated in RL, suggesting that they might enhance the red color accumulation, forming leucoanthocyanidin and transferring into anthocyanidins in *A. mandshuricum*. Previous studies have shown that the *ANS* gene is abundantly expressed in cell cultures of red cell lines and seedlings in the rare medicinal plant *Saussurea medusa*⁴⁴, and the high expression of *ANS* and *DFR* also promotes anthocyanin biosynthesis in red pear fruit skin⁴⁵; similar findings have been validated in *Litchi chinensis* Sonn⁴⁶ and *Malus hupehensis*³⁹. In addition, the anthocyanidin 3-O-glucosyltransferase (*BZ1*, EC: 2.4.1.115) is the last key enzyme in the anthocyanin biosynthetic pathway and can catalyze the transformation of unstable anthocyanidin into anthocyanin, including by catalyzing the transfer of glucose from UDP-glucose to anthocyanidins including delphinidin, cyanidin, etc.^{47–49}. As the leaves changed from green to red, two *BZ1* genes were down-regulated, two genes were up-regulated, and one gene was up-regulated only in half-red leaves. Notably, the *BZ1* gene *Cluster-20561.99238* (Table S8) was highly expressed, and its downstream metabolites (cyanidin 3-O-glucoside, pelargonidin 3-O-glucoside, and delphinidin 3-O-glucoside) were correspondingly significantly upregulated in the RL period (Fig. 6). Therefore, we suggested that *BZ1* gene was the key gene in regulating anthocyanin synthesis and leaf color stabilization. Interestingly, we found that the *UGT75C1* gene (anthocyanidin 3-O-glucoside 5-O-glucosyltransferase, EC:2.4.1.298) was significantly down-regulated in the anthocyanin synthesis process, leading to the formation of more derivatives of anthocyanins.

Some TFs, such as MYB, bHLH, and WD40, are key regulatory factors for structural genes in flavonoid biosynthesis pathways^{50,51}. The MYB transcription factor family is very large and versatile in plants, and these TFs can regulate the biosynthesis of phenylpropanoids⁵². Therefore, many MYB genes were identified, which were considered potential key regulators in the anthocyanin biosynthesis pathway⁵³. MYB, bHLH, and WD40 can be combined into the transcription factor complex MBW (MYB-bHLH-WD40) to regulate the biosynthesis of anthocyanins⁵⁴; they have been shown to regulate the synthesis of anthocyanins positively or negatively in plants including banana⁵⁵, rice⁵⁶, persimmon⁵⁷, etc. In the present study, a total of 341 (106 up-regulated and 235 down-regulated), 454 (103 up-regulated and 351 down-regulated) and 981 (346 up-regulated and 635 down-regulated) differentially expressed TFs were identified in GL vs HRL, HRL vs RL and GL vs RL, respectively (Table S9). Among them, we identified several candidate MYB (*Cluster-20561.80755*, *Cluster-20561.81808*, *Cluster-20561.96473*, and *Cluster-20561.98223*) and bHLH (*Cluster-20561.92650*) TFs that were highly up-regulated with most of the key genes in flavonoid biosynthesis during the discoloration of the leaves in *A. mandshuricum* and may positively regulate the key structural genes in anthocyanin synthesis in *A. mandshuricum* (Table S8). In addition, some TFs such as NAC, RWP-RK, WRKY, C2H2, GARP-G2-like, MYB-related, HD-Zip, and bZIP were discovered during the leaf color change periods and displayed different expression levels, indicating that

they may co-regulate the expression level of structural genes involved in the synthesis of flavonoids, similar to the results of recent studies^{58,59}.

Materials and methods

Plant materials. Fresh plant leaves of *A. mandshuricum* were harvested during three consecutive leaf color change periods, i.e., green leaf (GL), half red leaf (HRL) (During this period, the leaves began to change from green to red, and clearly visible red pigment began to appear on the leaves), and red leaf (RL) (Fig. 1). All samples were taken in the year 2020, green leaves were harvested at 10 a.m. on June 15, half red leaves were harvested at 10 a.m. on September 22, and red leaves were harvested at 10 a.m. on October 4. The leaves of all three periods were harvested from the same tree. The samples were frozen in liquid nitrogen and stored at -80°C until further use. Each color leaf contains three biological repeats. For each repeat, normal growth and no pests or diseased leaves were considered for one mixed sample³⁴, and the total sample weight was more than 3 g. These samples were used for subsequent transcriptome and metabolome studies.

Fresh samples were collected at the Northeast Forestry University (North: 45.723999; East: 126.635474) in Heilongjiang Province, China. And the experimental research on the plants described in this study comply with institutional, national and international guidelines.

Sample preparation and metabolite extraction. For metabolite extraction, five steps needed to be completed. (1) The samples were subjected to freeze-dry. (2) A mixer mill (MM 400, Retsch) was used to crush the freeze-dried samples with a Zirconia bead for 1.5 min at 30 Hz. (3) 50 mg powder was weighed and extracted with 0.5 ml methanol/water/hydrochloric acid (799:200:1, V/V/V). (4) The extractives were vortexed for 10 min, ultrasonicated for 10 min, and centrifuged at $12,000\times g$ under 4°C for 3 min, absorbed the supernatants. The step 4 repeated twice. (5) The supernatants were collected twice and filtered (PTFE, 0.22 μm ; Anpel) before LC-MS/MS (Liquid Chromatography-Tandem Mass Spectrometry, <https://sciex.com.cn>) analysis.

Metabolite analysis. Professional analytical processes of LC-MS/MS were performed by MetWare (<http://www.metware.cn/>), using the same methods described in previous articles^{60,61}. Metabolite data analysis was conducted with the R Programming language (<https://www.r-project.org/>), and hierarchical cluster analysis (HCA) and orthogonal partial least-squares discriminant analysis (OPLS-DA) were performed to identify the accumulation pattern in metabolites from nine leaf samples of three colors. The OPLS-DA model was used to identify the differences in metabolite composition between the samples by variable importance in projection (VIP) values ($\text{VIP} \geq 1$) and Fold change ≥ 2 or ≤ 0.5 .

RNA extraction, quantification and sequencing. RNA extraction, quantification and transcriptome sequencing were performed as fully described in Lu's study⁶². For transcriptome sequencing, we constructed nine libraries representing the three leaf stages with three biologically replicates for each period. All libraries were sequenced on an Illumina HiSeq platform, and raw data with 150 bp paired-end reads were obtained for subsequent analysis.

Transcriptome data analysis. The raw data (or raw reads) were filtered with fastq software (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) to filter out adapters and low-quality sequences and obtain high-quality clean sequences. Then, we used Trinity software (version 2.6.6) to perform De novo assembly. To obtain the annotation information, the unigenes were compared with public databases, including KEGG^{63,64}, GO, KOG, NR, Swiss-Prot, Pfam and Trembl. FPKM (fragments per kilobase of transcript per million fragments mapped) was used as an indicator to measure the level of transcript or gene expression. Differential expression analysis between different comparisons of colored samples was performed using the DESeq2 R package (1.20.0)⁶⁵. When the difference analysis was completed, the Benjamini-Hochberg method was used to carry out multiple hypothesis test corrections for the P value to obtain the false discovery rate (FDR). The screening conditions of unigenes were $|\log_2(\text{fold change})| \geq 1$ and $\text{FDR} < 0.05$ ⁶⁶. KEGG and GO enrichment analyses of differentially expressed genes were further implemented employing the KOBAS 2.0 software⁶⁷ and Goseq R packages⁶⁸, respectively.

Gene validation by real-time quantitative PCR (qRT-PCR). To further verify the reliability and accuracy of the transcriptome data, we initially screened 12 DEGs associated with phenylpropanoid, flavonoid, and anthocyanidin biosynthesis to detect expression levels in GL, HRL and RL with qRT-PCR. The Tiangen total RNA extraction kit (Tiangen, Beijing, China) and PrimeScript RT reagent Kit with gDNA Eraser (TaKaRa, Kyoto, Japan) were used for RNA extraction and cDNA synthesis, respectively. Primer 3 software (version 2.5.0, <https://primer3.ut.ee/>) was used to design specific primers for qRT-PCR analysis. qRT-PCR was carried out with the TaKaRa SYBR Green Mix kit (TaKaRa, Beijing, China) using an ABI 7500 Fast Real-Time Detection System. The PCR amplification experiment was performed with a total volume of 20 μl (10 μl 2xSYBR Primix ExTaq, 0.4 μl ROX-reference, 6 μl ddH₂O, 0.8 μl downstream primer, 0.8 μl upstream primer, and 2 μl cDNA template under a standard program (95°C for 30 s, 40 cycles at 95°C for 5 s, 60°C for 35 s, and 95°C for 15 s, and 60°C for 1 min, followed by 95°C for 15 s). The relative expression analysis of quantitative data was performed using the $2^{-\Delta\Delta\text{CT}}$ method⁶⁹ with reference gene 18S-RNA as a control.

Conclusion

In this study, the molecular mechanisms of anthocyanin biosynthesis in *A. mandshuricum* during different leaf color change periods were revealed based on metabolomics and transcriptomics studies. We analyzed the accumulation pattern of anthocyanins in three different leaf color states and found that cyanidin 3-O-glucoside was the predominant metabolite in the anthocyanin biosynthesis pathway, which accumulated in large quantities during the red leaf period. Furthermore, we analyzed the differentially expressed genes associated with anthocyanin biosynthesis in the leaves at different times. A variety of regulatory models of structural genes (*PAL*, *ANS*, *BZI*, etc.) and some potential transcription factors (*MYB*, *bHLH*, *C2H2*, *NAC*, *bZIP*, etc.) involved in biosynthesis of flavonoid pathways and anthocyanin biosynthesis pathway were found. The regulation of these key structural genes or their transcription factors might help elucidate the mechanism of leaf color change in *A. mandshuricum*. Our studies could provide novel insight into the molecular regulation of anthocyanin biosynthesis and accumulation associated with leaf color change in *A. mandshuricum*.

Data availability

Transcriptome sequencing data are available in the SRA database of National Center for Biotechnology Information (NCBI) under the accession number of PRJNA741917.

Received: 9 August 2021; Accepted: 18 November 2021

Published online: 30 November 2021

References

- Zhu, Y. Y. Research on relationships between functional traits and compensation mechanisms of androdioecious *Acer mandshuricum*, master thesis, Beijing Forestry University (2016).
- Liu, X. L. Research on red pigment in leaves and leaves color retention of *Acer mandshuricum*, master thesis, Northeast Forestry University (2014).
- Xie, Z., Yi, X. Y. & Xia, F. C. The resource situation and development utilization of *Acer* plants in northeast area. *J. Jilin For. Sci. Technol.* **46**, 26–31+47. <https://doi.org/10.16115/j.cnki.issn.1005-7129.2017.01.007> (2017).
- Reay, P. F., Fletcher, R. H. & Thomas, V. J. Chlorophylls, carotenoids and anthocyanin concentrations in the skin of “Gala” apples during maturation and the influence of foliar applications of nitrogen and magnesium. *J. Sci. Food Agric.* **76**, 63–71. [https://doi.org/10.1002/\(SICI\)1097-0010\(199801\)76:1](https://doi.org/10.1002/(SICI)1097-0010(199801)76:1) (2015).
- Englert, M., Hammann, S. & Vetter, W. Isolation of β -carotene, α -carotene and lutein from carrots by countercurrent chromatography with the solvent system modifier benzotrifluoride. *J. Chromatogr. A* **1388**, 119–125. <https://doi.org/10.1016/j.chroma.2015.02.020> (2015).
- Grotewold, E. The genetics and biochemistry of floral pigments. *Annu. Rev. Plant Biol.* **57**, 761–780. <https://doi.org/10.1146/annurev.arplant.57.032905.105248> (2006).
- Zhao, M. H., Li, X., Zhang, X. X., Zhang, H. & Zhao, X. Y. Mutation mechanism of leaf color in plants: A review. *Forests* **11**, 851. <https://doi.org/10.3390/f11080851> (2020).
- Rusishvili, M. *et al.* Unraveling the molecular mechanisms of color expression in anthocyanins. *Phys. Chem. Chem. Phys.* **21**, 8757–8766. <https://doi.org/10.1039/c9cp00747d> (2019).
- Tanaka, Y., Sasaki, N. & Ohmiya, A. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* **54**, 733–749. <https://doi.org/10.1111/j.1365-3113X.2008.03447.x> (2008).
- He, Q. *et al.* Natural variation in petal color in *Lycoris longituba* revealed by anthocyanin components. *PLoS ONE* **6**, e22098. <https://doi.org/10.1371/journal.pone.0022098> (2011).
- Martins, T. R., Berg, J. J., Blinks, S., Rausher, M. & Baum, D. A. Precise spatio-temporal regulation of the anthocyanin biosynthetic pathway leads to petal spot formation in *Clarkia gracilis* (*Onagraceae*). *New Phytol.* **197**, 958–969. <https://doi.org/10.1111/nph.12062> (2013).
- Sheng, T., Keurentjes, J., Bentsink, L. & Smeekens, K. S. Sucrose-specific induction of anthocyanin biosynthesis in *Arabidopsis* requires the MYB75/PAP1 gene. *Plant Physiol.* **139**, 1840–1852. <https://doi.org/10.1104/pp.105.066688> (2005).
- Wei, S. *et al.* Molecular and biochemical analysis of chalcone synthase from *Freesia* hybrid in flavonoid biosynthetic pathway. *PLoS ONE* **10**, e0119054. <https://doi.org/10.1371/journal.pone.0119054> (2015).
- Wei, H. *et al.* Comparative transcriptome analysis of genes involved in anthocyanin biosynthesis in the red and yellow fruits of *Sweet Cherry* (*Prunus avium* L.). *PLoS ONE* **10**, e0121164. <https://doi.org/10.1371/journal.pone.0121164> (2015).
- Watts, K. T., Lee, P. C. & Schmidt-Dannert, C. Exploring recombinant flavonoid biosynthesis in metabolically engineered *Escherichia coli*. *ChemBioChem* **5**, 500–507. <https://doi.org/10.1002/cbic.200300783> (2010).
- Liang, J. & He, J. Protective role of anthocyanins in plants under low nitrogen stress. *Biochem. Biophys. Res. Commun.* **498**, 946–953. <https://doi.org/10.1016/j.bbrc.2018.03.087> (2018).
- Huang, Y., Zhang, J., Fan, Y., Liu, Y. & Yang, M. Effects of shading treatment on leaf color and related physiological indexes of *Ulmus pumila* “Jinye” and *Koeleruteria paniculata* “Xinye”. *Scientia Silvae Sinicae* **55**, 171–180. <https://doi.org/10.11707/j.1001-7488.20191017> (2019).
- Kim, J., Woo, H. R. & Hong, G. N. Toward systems understanding of leaf senescence: an integrated multi-omics perspective on leaf senescence research. *Mol. Plant* **9**, 813–825. <https://doi.org/10.1016/j.molp.2016.04.017> (2016).
- Zhang, C., Wang, M., Gao, X., Zhou, F. & Liu, Z. Multi-omics research in albino tea plants: past, present, and future. *Sci. Hortic.* **261**, 108943. <https://doi.org/10.1016/j.scienta.2019.108943> (2020).
- Liu, X., Zhang, J., Luo, Y. & Rao, G. Metabolome and transcriptome analyses reveal tissue-specific variations in gene expression and metabolites of *Olive*. *J. Plant Biol.* **63**, 73–82. <https://doi.org/10.1007/s12374-020-09231-z> (2020).
- Sun, J. Transcriptome sequencing and metabolome analysis reveal genes involved in pigmentation of green-colored *Cotton* fibers. *Int. J. Mol. Sci.* **20**, 4838. <https://doi.org/10.3390/ijms20194838> (2019).
- Fang-Fang, F. U., Peng, Y. S., Wang, G. B., El-Kassaby, Y. A. & Cao, F. L. Integrative analysis of the metabolome and transcriptome reveals seed germination mechanism in *Punica granatum* L. *J. Integr. Agric.* **20**, 132–146. [https://doi.org/10.1016/S2095-3119\(20\)63399-8](https://doi.org/10.1016/S2095-3119(20)63399-8) (2021).
- Wan, L., Lei, Y., Yan, L., Liu, Y. & Liao, B. Transcriptome and metabolome reveal redirection of flavonoids in a white testa peanut mutant. *BMC Plant Biol.* **20**, 1–19. <https://doi.org/10.1186/s12870-020-02383-7> (2020).
- Qiu, W., Su, W., Cai, Z., Dong, L. & Wu, Z. Combined analysis of transcriptome and metabolome reveals the potential mechanism of coloration and fruit quality in yellow and purple *Passiflora edulis* Sims. *J. Agric. Food Chem.* **68**, 12096–12106. <https://doi.org/10.1021/acs.jafc.0c03619> (2020).

25. Wei, Y., Zhang, Y., Meng, J., Wang, Y. & Zhong, C. Transcriptome and metabolome profiling in naturally infested *Casuarina equisetifolia* clones by *Ralstonia solanacearum*. *Genomics* **113**, 1906–1918. <https://doi.org/10.1016/j.ygeno.2021.03.022> (2021).
26. Altschul, S. F. Basic local alignment search tool (BLAST). *J. Mol. Biol.* **215**, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) (2012).
27. Varet, H., Brillet-Guéguen, L., Coppée, J. & Dillies, M. A. SARTools: a DESeq2- and EdgeR-Based R Pipeline for comprehensive differential analysis of RNA-Seq data. *PLoS ONE* **11**, e0157022. <https://doi.org/10.1371/journal.pone.0157022> (2016).
28. Yamagishi, M. Involvement of a LhMYB18 transcription factor in large anthocyanin spot formation on the flower tepals of the Asiatic hybrid lily (*Lilium* spp.) cultivar “Grand Cru”. *Mol. Breed.* **38**, 60. <https://doi.org/10.1007/s11032-018-0806-1> (2018).
29. Filiz, E. & Kurt, F. Expression and Co-expression analyses of WRKY, MYB, bHLH and bZIP transcription factor genes in Potato (*Solanum tuberosum*) under abiotic stress conditions: RNA-seq data analysis. *Potato Res.* <https://doi.org/10.1007/s11540-021-09502-3> (2021).
30. Pireyre, M. & Burow, M. Regulation of MYB and bHLH transcription factors: a glance at the protein level. *Mol. Plant* **8**, 378–388. <https://doi.org/10.1016/j.molp.2014.11.022> (2015).
31. Zongyan, L. I. & Mingyang, L. I. Advance in transcriptional factors regulating flavonoid biosynthesis. *J. Nanjing For. Univ.* **35**, 129–134. <https://doi.org/10.1631/jzus.b1000171> (2011).
32. Duan, S. *et al.* Functional characterization of a heterologously expressed *Brassica napus* WRKY41-1 transcription factor in regulating anthocyanin biosynthesis in *Arabidopsis thaliana*. *Plant Sci.* **268**, 47–53. <https://doi.org/10.1016/j.plantsci.2017.12.010> (2018).
33. Li, X., Xiang, F., Han, W., Qie, B. & Xu, L. The MIR-Domain of PbbHLH2 is involved in regulation of the anthocyanin biosynthetic pathway in “Red Zaosu” (*Pyrus bretschneideri* Rehd.) pear fruit. *Int. J. Mol. Sci.* **22**, 3026. <https://doi.org/10.3390/ijms22063026> (2021).
34. Zhang, Q., Wang, L., Liu, Z., Zhao, Z. & Liu, M. Transcriptome and metabolome profiling unveil the mechanisms of *Ziziphus jujuba* Mill. peel coloration. *Food Chem.* **312**, 125903. <https://doi.org/10.1016/j.foodchem.2019.125903> (2019).
35. Chen, Z. *et al.* Transcriptome analysis based on a combination of sequencing platforms provides insights into leaf pigmentation in *Acer rubrum*. *BMC Plant Biol.* **19**, 240. <https://doi.org/10.1186/s12870-019-1850-7> (2019).
36. Becker, C., Klaering, H. P., Kroh, L. W. & Krumber, A. Cool-cultivated red leaf lettuce accumulates cyanidin-3-O-(6"-O-malonyl)-glucoside and caffeoylmalic acid. *Food Chem.* **146**, 404–411. <https://doi.org/10.1016/j.foodchem.2013.09.061> (2014).
37. Gao, Y. F. *et al.* De novo transcriptome sequencing and anthocyanin metabolite analysis reveals leaf color of *Acer pseudosieboldianum* in autumn. *BMC Genomics* **22**, 383. <https://doi.org/10.1186/s12864-021-07715-x> (2021).
38. Wang, X., Bai, J., Wang, W. & Zhang, G. Leaf metabolites profiling between red and green phenotypes of *Suaeda salsa* by widely targeted metabolomics. *Funct. Plant Biol.* **46**, 845–856. <https://doi.org/10.1071/FP18182> (2019).
39. Han, M., Yang, C., Zhou, J., Zhu, J. & Li, H. Analysis of flavonoids and anthocyanin biosynthesis-related genes expression reveals the mechanism of petal color fading of *Malus hupehensis* (Rosaceae). *Braz. J. Bot.* **43**, 81–89. <https://doi.org/10.1007/s40415-020-00590-y> (2020).
40. Li, W. *et al.* The effects of ultraviolet A/B treatments on anthocyanin accumulation and gene expression in dark-purple tea cultivar “Ziyan” (*Camellia sinensis*). *Molecules* **25**, 354. <https://doi.org/10.3390/molecules25020354> (2020).
41. Zhang, Y. Z., Xu, S. Z., Cheng, Y. W., Ya, H. Y. & Han, J. M. Transcriptome analysis and anthocyanin-related genes in red leaf lettuce. *Genet. Mol. Res.* **15**, 15017023. <https://doi.org/10.4238/gmr.15017023> (2016).
42. He, H. *et al.* Flower colour modification of *Chrysanthemum* by suppression of F3'H and overexpression of the exogenous *Senecio cruentus* F3'5'H gene. *PLoS ONE* **8**, e74395. <https://doi.org/10.1371/journal.pone.0074395> (2013).
43. Han, Y. *et al.* Introduction of apple ANR genes into tobacco inhibits expression of both CHI and DFR genes in flowers, leading to loss of anthocyanin. *J. Exp. Bot.* **63**, 2437–2447. <https://doi.org/10.1093/jxb/err415> (2012).
44. Cheng, L. *et al.* Characterization of Anthocyanidin Synthase (ANS) Gene and anthocyanidin in rare medicinal plant-*Saussurea medusa*. *Plant Cell Tissue Organ Cult.* **89**, 63–73. <https://doi.org/10.1007/s11240-007-9211-x> (2007).
45. Li, X., Wu, T., Liu, H., Zhai, R. & Xu, L. Reveille transcription factors contribute to the nighttime accumulation of anthocyanins in “Red Zaosu” (*Pyrus bretschneideri* Rehd.) pear fruit skin. *Int. J. Mol. Sci.* **21**, 1634. <https://doi.org/10.3390/ijms21051634> (2020).
46. Wei, Y. Z. *et al.* Differential expression of anthocyanin biosynthetic genes in relation to anthocyanin accumulation in the pericarp of *Litchi Chinensis* Sonn. *PLoS ONE* **6**, e19455. <https://doi.org/10.1371/journal.pone.0019455> (2011).
47. Hiromoto, T. *et al.* Structural basis for acceptor-substrate recognition of UDP-glucose: Anthocyanidin 3-O-glucosyltransferase from *Clitoria ternatea*. *Protein Sci.* **24**, 395–407. <https://doi.org/10.1002/pro.2630> (2015).
48. Du, L. J., Chen, K. L. & Liu, Y. L. Cloning and expression analysis of anthocyanidin 3-O-glucosyltransferase gene in Grape hyacinth. *Pratacultural Sci.* **34**, 2235–2244. <https://doi.org/10.11829/j.issn.1001-0629.2017-0049> (2017).
49. Hongxia, *et al.* UDP-glucose: Anthocyanidin 3-O-glucoside-2"-O-glucosyltransferase catalyzes glycosyl extension of anthocyanins in purple *Ipomoea batatas*. *J. Exp. Bot.* **69**, 5445–5459. <https://doi.org/10.1093/jxb/ery305> (2018).
50. Cui, X., Deng, J., Huang, C., Tang, X. & Zhang, Z. Transcriptomic analysis of the anthocyanin biosynthetic pathway reveals the molecular mechanism associated with purple color formation in *Dendrobium Nestor*. *Life* **11**, 113. <https://doi.org/10.3390/life11020113> (2021).
51. Yao, Y., Sun, L., Wu, W. J., Wang, S. & Wu, Q. Genome-wide investigation of major enzyme-encoding genes in the flavonoid metabolic pathway in *Tartary Buckwheat* (*Fagopyrum tataricum*). *J. Mol. Evol.* **89**, 269–286. <https://doi.org/10.1007/s00239-021-10004-6> (2021).
52. Arce-Rodríguez, M., Martínez, O. & Ochoa-Alejo, N. Genome-wide identification and analysis of the MYB transcription factor gene family in *Chili Pepper* (*Capsicum* spp.). *Int. J. Mol. Sci.* **22**, 2229. <https://doi.org/10.3390/ijms22052229> (2021).
53. Zhu, Y. & Bao, Y. Genome-wide mining of MYB transcription factors in the anthocyanin biosynthesis pathway of *Gossypium hirsutum*. *Biochem. Genet.* **59**, 678–696. <https://doi.org/10.1007/s10528-021-10027-0> (2021).
54. Liu, Y. *et al.* Comparative transcriptome analysis of white and purple Potato to identify genes involved in anthocyanin biosynthesis. *PLoS ONE* **10**, e0129148. <https://doi.org/10.1371/journal.pone.0129148> (2015).
55. Deng, G. M., Zhang, S., Yang, Q. S., Gao, H. J. & Hu, C. H. MaMYB4, an R2R3-MYB repressor transcription factor, negatively regulates the biosynthesis of anthocyanin in Banana. *Front. Plant Sci.* **11**, 600704. <https://doi.org/10.3389/fpls.2020.600704> (2021).
56. Qiao, W. H. *et al.* A functional chromogen gene C from wild rice is involved in a different anthocyanin biosynthesis pathway in indica and japonica. *Theor. Appl. Genet.* **134**, 1531–1543. <https://doi.org/10.1007/s00122-021-03787-1> (2021).
57. Gil-Muoz, F., Sánchez-Navarro, J., Besada, C., Salvador, A. & Ríos, G. MBW complexes impinge on anthocyanidin reductase gene regulation for proanthocyanidin biosynthesis in persimmon fruit. *Sci. Rep.* **10**, 3543. <https://doi.org/10.1038/s41598-020-60635-w> (2020).
58. Ahmad, S. *et al.* Transcriptional cascade in the regulation of flowering in the Bamboo Orchid *Arundina graminifolia*. *Biomolecules* **11**, 771. <https://doi.org/10.3390/biom11060771> (2021).
59. Premathilake, A. T., Ni, J., Shen, J., Bai, S. & Teng, Y. Transcriptome analysis provides new insights into the transcriptional regulation of methyl jasmonate-induced flavonoid biosynthesis in pear calli. *BMC Plant Biol.* **20**, 388. <https://doi.org/10.1186/s12870-020-02606-x> (2020).
60. Qin, G., Liu, C., Li, J., Qi, Y. & Xu, Y. Diversity of metabolite accumulation patterns in inner and outer seed coats of pomegranate: exploring their relationship with genetic mechanisms of seed coat development. *Horticult. Res.* **7**, 10. <https://doi.org/10.1038/s41438-019-0233-4> (2020).

61. Chen, Z., Lu, X., Gao, J., Xuan, Y. & Ren, J. Integrating transcriptomic and metabolomic analysis of hormone pathways in *Acer rubrum* during developmental leaf senescence. *BMC Plant Biol.* **20**, 410. <https://doi.org/10.1186/s12870-020-02628-5> (2020).
62. Lu, D. *et al.* Transcriptome analysis of chrysanthemum in responses to white rust. *Sci. Hortic.* **233**, 421–430. <https://doi.org/10.1016/j.scienta.2018.01.016> (2018).
63. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951. <https://doi.org/10.1002/pro.3715> (2019).
64. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. & Tanabe, M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**, D545–D551. <https://doi.org/10.1093/nar/gkaa970> (2021).
65. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8> (2014).
66. Shi, Q. Q., Du, J. T., Zhu, D. J., Li, X. & Li, X. G. Metabolomic and transcriptomic analyses of anthocyanin biosynthesis mechanisms in the color mutant *Ziziphus jujuba* cv. Tailihong. *J. Agric. Food Chem.* **68**, 15186–15198. <https://doi.org/10.1021/acs.jafc.0c05334> (2020).
67. Chen, X. *et al.* KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* **39**, 316–322. <https://doi.org/10.1093/nar/gkr483> (2011).
68. Young, M., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **11**, R14. <https://doi.org/10.1186/gb-2010-11-2-r14> (2010).
69. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR. *Methods* **25**, 402–408. <https://doi.org/10.1006/meth.2001.1262> (2002).

Acknowledgements

This work was supported by the Heilongjiang Province Applied Technology Research and Development Plan Project (No. GA20B402), Fundamental Research Funds for the Central Universities (Northeast Forestry University) (No. 2572020DR01) and Heilongjiang Touyan Innovation Team Program (Tree Genetics and Breeding Innovation Team).

Author contributions

X.Z., S.Z. and W.Z. conceived the experiment(s) and wrote the manuscript, A.S., Y.X., Z.H., X.Q. and Y.J. conducted the experiment(s) and analysed the results. J.Z. and M.T. consulted literatures and revised the grammar, L.Z. and X.P. made the final revision to the article. All authors reviewed the manuscript and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-02607-2>.

Correspondence and requests for materials should be addressed to X.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021